*** **DRAFT** ***

Modeling Protocols for Water & Environmental Problem-Solving

Bay-Delta Modeling Forum
Ad hoc Modeling Protocols Committee

February 24, 1998
(Slightly Revised on March 6, 1998)

Foreword

This draft report presents modeling protocols that provide the water community with basic principles and guidelines for model development and use. This effort was spawned by discussion during the Bay-Delta Modeling Forum's (Forum) 1996 Annual Meeting & Workshop breakout session entitled "Should or Can Modeling Protocols/Conventions be Standardized?" At that session, the participants hypothesized that water stakeholders and decision-makers often lose confidence in models because of inconsistencies in the way models are developed and used. After much discussion, the participants concluded that uniform application of modeling protocols should result in better models and modeling studies, and, thus, increase the confidence of stakeholders and decision-makers who use model results. Consequently, the breakout participants unanimously agreed that modeling protocols can and should be standardized, and that the Forum should take the lead in this effort.

In March of 1997, the Forum formed an Ad hoc Modeling Protocols Committee to (1) develop modeling protocols that can become standards for model development and use and (2) prepare a written report of findings for Forum acceptance. As part of this effort, the Modeling Protocols Committee developed the following mission statement:

> The mission of the Modeling Protocols Committee is to develop modeling principles and guidelines (protocols) that provide guidance to water stakeholders and decision-makers, and their technical staff as models are developed and used to solve California's water and environmental problems.

The Ad hoc Modeling Protocols Committee (listed below) expects to complete this report by May 1998. If the Forum "accepts" the final report, the committee will assist Forum members and other interested parties in implementing the modeling protocol recommendations contained in this report. In addition, the committee will develop additional modeling protocols, if necessary. As specified in Section 9.05 of the Forum bylaws, it should be noted that this report does not necessarily represent the views of the governing bodies of the represented organizations or individual members of the Forum.

<u>Ad hoc Modeling Protocols Committee</u>

Richard Satkowski (SWRCB)-Chair *

Betty Andrews (PWA)	Hubert Morel-Seytoux (HDP) *
George Barnes (DWR)	Austin Nelson (Consultant) *
Sam Bledsoe (UCD)	Ted Roefs (USBR) *
Grace Chan (MWD)	Spreck Rosekrans (EDF)
Wim Kimmerer (RTC)	Jim Sutton (SWRCB)
Jay Lund (UCD) *	Peter Vorster (TBI)

John Williams-Executive Director

Modeling Protocols for Water and Environmental Problem-Solving

Contents

1. Introduction

Water stakeholders and decision-makers use models as a tool to help solve California's water and environmental problems. Unfortunately, stakeholders and decision-makers, and even their technical staff, often lose confidence in models because of (1) an inadequate understanding about modeling and (2) inconsistencies in the way models are developed and used. Confidence in model predictions and a good understanding of modeling is essential for stakeholders and decision-makers responsible for setting water quality standards, flow requirements and other regulations. To address this problem, the Bay-Delta Modeling Forum (Forum) has developed modeling protocols, which are basic principles and guidelines for model development and use. Model developers, users of modeling services, and water stakeholders and decision-makers wishing to understand modeling and its consequences should use modeling protocols.

The objective of these modeling protocols are to provide guidance to water stakeholders and decision-makers, and their technical staff as models are developed and used to solve California's water and environmental problems. The Forum believes that acceptance and implementation of modeling protocols by California's water community will result in better models and modeling studies by doing the following:

- Improving the construction of models;
- Providing better documentation of models and modeling studies;
- Providing easier public access to models and modeling studies;
- Making models and modeling studies more easily understood and more amendable to examination; and
- Increasing stakeholder, decision-maker, and technical staff confidence in models and modeling studies.

A computer model consists of two basic parts: the computer code or software and the input data set. Computer models can be as simple as a mass balance equation, which can be performed on a calculator, to multiple differential equations that require high-speed computers. According to the Forum bylaws, modeling includes, but is not limited to, the following water-related topics (BDMF, 1997):

- Data gathering, storage and access
- Economics
- Fisheries, aquatic biology, and habitat health
- Groundwater
- Hydrodynamics
- Hydrology, hydraulics, and irrigation
- System operations and real-time management
- Water quality
- Water resources planning

This report explains in basic terms why models are important, how modeling efforts are reviewed, and how models are developed and used. In addition, it describes various modeling protocols that provide a consistent framework to develop, apply, and document a computer model.

1. Purposes of Modeling

Solving California's Water Problems

Computer models do not resolve conflicts: people do. However, computer modeling can assist in that role by doing the following (Lund and Palmer, 1998):

- Furthering understanding of the problem.
- Defining solution objectives.
- Developing promising alternatives.
- Evaluating alternatives.
- Providing confidence in solutions.
- Providing a forum for negotiations.

The purpose of a model is to reproduce consistently the observable phenomena that are of significance for a particular problem. For example, the purpose of a salinity water quality model is to reproduce in time and space the distribution of salinity due to the effects of flows, diversions, tides, etc. Modeling can be used to support real-time decision-making or evaluate a physical or biological system under historical, present and future conditions.

California resources planning is increasingly dependent on analytical methods and tools (models) that can provide practical answers for immediate problems and significant direction for long-range plans (BDMF, 1995). Models are essential tools for analysis of issues arising in water rights and development of new projects. These models play an important role in developing environmental impact analyses of projects under the California Environmental Quality Act (CEQA), and the federal National Environmental Policy Act (NEPA), such as the CALFED Bay-Delta Process, Central Valley Project Improvement Act (CVPIA), Interim South Delta Program, Contra Costa Water District's Los Vaqueros Project, the Delta Wetlands Project, and other projects.

Historical Solutions (Under Development)

Introduction

Water Supply Planning - the California Water Plan

Bulletin 27 Model
USCE Physical Model

Water Supply Planning - State Water Project

Seven-Reach Model - Salt Routing Model and Carriage Water
DAYFLOW - Data Compilation and Analysis - QWEST

Delta Facilities Planning and the Peripheral Canal

PCSTAGE - Evaluating Consequences of Project Staging
FLOSALT - Delta Agricultural Drainage Assessments

San Francisco Bay-Delta Water Quality Control Plan

Link-Node Models
TVSALT - Extension of the Link-Node Models
Striped Bass Model

San Luis Drain

Fischer Delta Model and its derivatives

Water Operations Models

State Water Project - DWRSIM
Central Valley Project - PROSIM
Los Vaqueros Project

Bay-Delta Water Quality Control Plan

Kimmerer-Monismith Equation
G-Model
Long Fin Smelt Model
Stary Flounder Model
Crangon Model
Salmon Survival Model
DELCORN
Year Type Models

Closing Notes

3. Stakeholder and Public Review of Modeling Efforts

General Public Participation

The planning process, or as it is sometimes called "decision-making," are the actions that lead to selection of a recommended plan. Planning should include an early and open process, termed "scoping," to identify both the likely significant issues to be addressed and the range of those issues. Scoping should be used throughout planning to ensure that all significant factors are addressed. Scoping may be used to narrow the number of plans under consideration so that meaningful and efficient analysis and choice among alternative plans can occur. Scoping should include consideration of all water problems and opportunities, including instream flow problems and conjunctive use of surface and ground water. Appropriate consideration should be given to existing water rights in scoping the planning effort.

Proper planning requires adequate review and consultation with interested and affected stakeholders, agencies, organizations, and individuals. These groups and individuals should be provided opportunities to participate throughout the planning process. Efforts to secure public participation should be pursued through public workshop, meetings, and technical advisory and citizens committees. (U.S. Water Resource Council, 1983)

Much of the failure of water projects is due to the lack of stakeholder and decision-maker communication as well as public participation in the water planning process. Failures in the process are usually due to faulty assumptions, faulty predictions of likely behavior of affected parties, and faulty assumptions of affected party perceptions. To overcome these failures, the planning process must be performed with as much dialog, input and agreement (consensus) as possible.

Shared Vision Modeling

In the technical and political arenas of California water, it is important that models enjoy a wide base of support from stakeholders and decision-makers, and their technical staff. Shared vision modeling is the common development of a model by a group of stakeholders and/or decision-makers. The fundamental concept is that those that will be impacted by water resource modeling should be provided the opportunity to participate in model design, development, evaluation, and enhancement. A goal of this process is to provide all interested parties with a tool increase their understanding of the problem and possible solutions. (Lund and Palmer, 1998). This approach is really an extension of classical engineering planning to more pluralistic decision-making circumstances (Werick and Whipple, 1994). The model is typically developed by a single, often neutral, entity with very close coordination by technical representatives from each stakeholder or stakeholder group. The model is then approved by the participants and can be used separately by each group, with a fixed model version and documentation (Lund and Palmer, 1998).

Shared vision modeling is intended to take the mostly technical decisions out from the political spotlight, and remove as many technical questions disagreements as possible from the conflict. If participants can arrive at agreement on what is contained in the model, then later efforts can focus on interpretation of the results, rather than arguments about model content. The process of developing this model is usually seen as a prelude to the developing and evaluating alternatives and meaningful negations among stakeholders. In addition, this approach helps to create a technically based forum where the parties can negotiate (Lund and Palmer, 1998).

Shared vision modeling, like other consensus building processes, requires that strong motivation exists among the stakeholders to develop a consensus. Arriving at a consensus about a model is not easy. Model development will progress much more slowly than if performed by single group (Lund and Palmer, 1998). However, if one considers the modeling and negotiation steps as one extended process, shared vision modeling usually saves time in the long run.

Peer-Review

The Bay-Delta Modeling Forum has developed a peer review process for peer reviewing computer models process (BDMF, 1996). Peer review is a method for reviewing models in a timely, open, fair, and helpful manner. Peer review serves two principle clients: model developers and model users by (1) providing constructive feedback to model developers and (2) serving to further the models' acceptance and understanding by the user community, including stakeholders and decision-makers.

These peer reviews are not intended to be "stamps-of-approval" for particular models or to disapprove of models. Instead, it is intended to inform stakeholders and decision-makers of (1) whether or not a given model is a suitable tool, and (2) the temporal, geographic, or other limits on the use of the model. The Forum's model peer review steps are as follows:

1. Select Models
2. Select Reviewers
3. Obtain Funding
4. Assemble Model, Documentation, and Data
5. Scope the Review
6. Conduct Initial Review
7. Test Models
8. Prepare Draft Report
9. Conduct Review Workshops
10. Prepare Final Report

For more information on the Forum's peer review process, access the Forum's webpage at *www.sfei.org/modelingforum/*.

External Review

California has enacted a "peer review" requirement for technical analysis performed by the California Environmental Protection Agency (Cal EPA). This law, Senate Bill 1320 (Sher), requires all organizations within the Cal EPA, such as the State Water Resources Control Board, to conduct an external scientific peer review of the scientific basis for any rule and prescribe procedures for conducting that scientific peer review (California Senate, 1997).

Under this law, the organizations within the Cal EPA can enter into an agreement with one or more of the following:

D-010086

- National Academy of Sciences;
- University of California;
- California State University;
- Any similar institution of higher learning; or
- A scientist or group of scientists of comparable stature and qualifications that is recommended by the President of the University of California.

Given the diverse and highly technical composition of the Bay-Delta Modeling Forum, it is likely that the Forum would qualify under the last category above. If the state organization disagrees with any aspect of the external scientific peer review, it must (1) explain why it disagrees in the adoption of the final rule, and (2) include this information as part of the rulemaking record. Senate Bill 1320 can be found on the Internet at: *www.leginfo.ca.gov.*

4. Model Development

> "Different types of models are appropriate for solving different kinds of problems; there is no universal model for solving all manner of problems; comprehensiveness and complexity in a simulation are no longer equated with accuracy; and there is a healthy mood of critical questioning of the validity and credibility of water quality models." M.B. Beck (1985)
>
> "All mathematical models are based on a set of simplifying assumptions, that affect their use for certain problems. ... To avoid applying an otherwise valid model to an inappropriate field situation, knowledge of all of the assumptions that form the basis of the model and consideration of their applicability to the site and problem under evaluations is very important." ASTM (1995a)

Computer models represent a systematic organization of our knowledge of a system developed for some planning, engineering, or scientific purpose. This chapter is divided into two sections. The first section discusses how different forms of knowledge are represented in computer models. The second section presents a relatively accepted approach for computer model development that (1) emphasizes the use of a model for problem-solving and (2) informs modelers and model users of general model strengths and limitations.

Knowledge Basis for Model Development

Models represent existing or hypothesized knowledge of how a system works. There are two major origins of this knowledge, causal (based on fundamentals of physics, chemistry, etc.) and empirical (based more directly on field or laboratory observations). These two bases for modeling are discussed frequently for water resource and environmental management models (Beck 1983a, 1985; Klemes 1982; Scavia and Chapra 1977). Since our knowledge of these systems is imperfect, probability is sometimes used in modeling to represent uncertainty. Also, models are often discussed as being vital to adaptive management (Holling 1978), where our knowledge of the system evolves with our management of it.

Mechanistic Models

Often called causally based or physically based models, mechanistic models rely on the fundamental rules of logic and laws of physics, chemistry, etc. Some examples of mechanistic models include the following:

- Use of conservation of mass to derive models of the operation of river-reservoir systems;
- Use of conservation of mass, momentum, and energy with channel geometries and bed elevations for hydraulic routing; and
- Use of principles of advection and dispersion for contaminant transport modeling.

Mechanistic models of physical phenomena commonly consist of a set of fundamental governing equations representing conservation of mass, energy, and momentum, reaction kinetics, etc. Often, these are differential equations. These governing equations have initial or boundary conditions, and can be solved by several numerical schemes.

Establishment of boundary conditions often requires a great deal of empirical knowledge, often with detailed spatial and temporal resolution. Also, many components of the governing equations are often empirical, such as use of Manning's equation for bed friction in hydraulic models and Fickian diffusion to represent dispersion in fate and transport models. The numerical solution techniques used to solve the large number of governing equations also involve some simplification of the system, in terms of simplification of the system's geometry and dynamics, which can induce some errors in model results. Thus, it is difficult to have a purely mechanistic model.

Empirical Models

The equations and calibrations of empirically-based models rely more directly on field or laboratory data, or empirical observations. Physical, chemical, biological, or socio-economic theory are of less importance than the accumulation of observations and data.

Empirical models sometimes can "fit" current experiences well, but are thought to be less reliable when the system is changed significantly, so that conditions change substantially from the behavior for which the model was developed (Klemes 1982).

Mixed Models

Most models used in water resources and environmental problem solving are mixtures of mechanistic and empirical models. The better-understood parts of the modeling problem (such as conservation of mass) are commonly mechanistic, whereas less well-understood processes, such as fluid friction are modeled based on empirical relationships (such as Manning's n).

Conceptual models are a common compromise between causally based and empirical models. Conceptual models often begin as a rough, relatively qualitative representation of how components of a model interact, based on theoretical, empirical, or hypothetical relationships. These "models" then can develop into quasi-mechanistic, quasi-empirical models. The Stanford Watershed Model of the 1970s is a fairly successful example of such a conceptual model.

Probabilistic Models

Analytical vs. Monte Carlo modeling (Under Development)

Adaptive Management and Computer Models

Much has been written about the role of computer models for improving the basis for management, even where little is known about the system being managed (Holling 1978). In these cases, computer models often are seen as a rigorous tool for systematically identifying what is known and what is not, estimating the importance of things imperfectly known, and providing an explicit technical basis for beginning to manage critical systems before they are completely understood. ... Models are both a distillate of past experience and a stimulus to the future development of experience (Beck 1985).

To improve, substantiate, and communicate the knowledge-based representation of a system in a model, model development usually proceeds along a series of commonly accepted steps.

Model Development Process

Considerable consensus exists in the profession and academia that the development of computer models should follow a somewhat standard procedure, which is outlined below and summarized in Table 1. This process is designed to aid the user of the model and the users of model results by providing assurances that the model works and identifying the limits of the model's capabilities. Additional discussion of this process can be found elsewhere (ASTM 1992, 1995a,b; Beck 1983a, 1985; Gass and Thompson 1980; James 1993; Jacoby and Kowalik 1980; Sargent 1988).

Table 1: Major Steps in Model Development

Step	Name	Purpose
1.	Problem Identification	Solving the right problem
2.	Define Modeling Objectives	Define use for model and standard of success
3.	Formulation of Model	Mathematical similarity to the problem system
4.	Selection and study of numerical solution	Numerical similarity to the mathematical formulation of the problem
5.	Model Calibration	Set constants to represent system behavior and characteristics
6.	Model Verification	Test model based on model behavior
7.	Model Validation	Test model by comparison with field data
8.	Documentation of Model	Make model understandable to users
9.	Update and Support of Model	Maintain and improve the model's usefulness

Although there is a fair consensus on the general contents of good model development, this procedure is not standardized to the degree found in many other technical fields, such as chemical analysis. Perhaps this lack of detailed standards reflects the difficulty and diversity of modeling problems. It also should be noted that the procedures discussed here are often iterative in nature. For example, failure of a model to calibrate well often leads to reexamination of the model's formulation. Also, detailed and prolonged work in model testing can identify new modeling objectives. As can be seen for most models commonly used in practical problem-solving, model development is rarely a completed process, but rather one of continual improvement, adaptation, and updating.

Step 1. Problem Identification

It is impossible to model everything. The first and arguably the most important step in modeling is to identify the problem to be modeled, and by implication identifying the problems (or parts of the problem) that are not to be addressed. Considerable attention should be given to the roles the model is expected to serve in addressing the problem, both in the short and long terms. Who will be using the model, and how is the model expected to help?

Step 2. Define Modeling Objectives

Most computer modeling efforts can address only a few important aspects of a general problem. Thus, it is important early on to identify specific modeling objectives. These objectives help the model development process by doing the following:

- Allowing the developers to focus on particular aspects of the problem and uses of the model;

- Providing a set of criteria for evaluating the performance of the model (i.e., how well does the model's application satisfy the stated objectives?); and

- Providing clear indications of intended model uses for potential users of the model and model results.

Few modeling decisions exist that should not be made without consideration of the objectives of the model and its problem context. Thus, the modeling objectives should also reflect some clear understandings of how the model is expected to be integrated into larger decision-making, scientific, or engineering problem-solving contexts.

Step 3. Formulation of a Model

Model formulation is the simplification of an understanding of the real problem to a mathematical form in a way consistent with the modeling objectives. Formulation involves the explicit specification of relationships thought to govern the behavior of the system (Beck 1983a).

Model formulation typically begins with development of a conceptual model, a working understanding of how a system works. This conceptual model usually represents our theoretical understanding of the system. This conceptual model forms the basis for the more detailed and explicit development of a mathematical model, a system of equations, typically implemented on a computer.

D-010091

In formulating a model, various decisions will need to be made, hopefully reflecting the problem, modeling objectives, and our understanding of the problem. In addition to formalizing the relationships that describe the system, the spatial and temporal aggregation and scale of the system need to be specified. Is the model to be steady-state or dynamic? Linear or non-linear? Deterministic or stochastic? If stochastic, which type of stochastic?

For example, for ecological modeling problems, the modeler needs to know which species of classes of species will be represented and which environmental factors affecting them are to be included? For hydrodynamic models, should 1, 2, or 3 dimensional representations of the system be used, how coarse a spatial grid, and if the model is dynamic, what time-step should be used? For water quality models, which constituents should be included, and how should their sources, sinks, and reactions be represented?

The purist's decision in all these cases is to choose the more detailed solution (highly disaggregated in time and space, dynamic, stochastic, etc.). However, this is usually the wrong decision, or perhaps merely an impossible decision. Highly detailed formulations are typically unsupported by field or laboratory data of sufficient quality or quantity and may provide little predictive understanding for the problem at hand. Complex models also are not always needed for the problem-solving objective. Instead, simplification commonly is needed, attempting to represent the most important parts of the problem, consistent with data and knowledge available within the context of the problem. For problem solving, we often cannot wait for perfect knowledge of a system. Indeed, the formulation and testing of a model (as an active hypothesis) usually can accelerate our understanding of a system. As a practical matter, there is need for a rough balance between the errors from simplification and the errors introduced by having additional uncertain parameters, inputs, and boundary conditions (Beck 1985).

As model development proceeds to model calibration, testing, and use, it is often necessary to revisit decisions made in model formulation, at least in part. Model development is usually an iterative process. This is healthy.

Step 4. Selection and study of numerical solution behavior.

Once the mathematical form of the model has been specified, the solution method for the model equations must be found. Often, particularly for complex models, the solution method for the model equations will require testing to ensure that the numerical solutions are correct for the intended types of problems and modeling objectives. Sometimes, concerns over numerical solution are reduced or eliminated through the use of commonly accepted software capable of solving some common forms of mathematical equations. These commonly accepted software can include spreadsheets, commercial equation solvers (e.g., MATLAB, LINDO, MINOS, etc.), or commercial subroutines (e.g., IBM's IMSL routines).

Step 5. Model Calibration.

Model calibration is the process of establishing specific values for parameters (constants) in the model's mathematical equations and algorithms. Typically, the purpose of calibration is to "fit" the model to the system being modeled, trying to "match" model and real output. The definition of a "good" fit or match between model and real output usually depends on the objectives and intended uses of the model. For example, for a rainfall-runoff model, if flood periods are of greatest interest, the ability of the model to correctly predict low flows might not be important.

Table 2 briefly describes several approaches to model calibration. Each method requires successively greater amounts of data from the real system.

Table 2: Approaches to Model Calibration

Approach	**Description**
Classical Physical Constants	Usually, there are physical constants, such as gravitational acceleration w known constant values. These parameters are typically set to these know values.
Literature Values	Often specific studies have been conducted elsewhere or at other times to estimate the value of parameters in specific model equations. These valu are often useful for estimating reasonable parameter values for other mod utilizing the same model equation in similar conditions. A variant of this approach is to have an "expert" on a particular parameter give an educate guess of what its value should be. This calibration approach is often used where data collection is impossible or to see if parameter values given by other approaches are "reasonable."
Field Measurements	Some model parameters, such as watershed area, are relatively determinis unchanging, and easily estimated. Field measurement or map measureme of such parameters can often give reasonable estimates.
Statistical	Very frequently, a model parameter can be measured, but might not have constant value. This can arise because there may be measurement error o natural variation of the parameter over time or space. If a single paramet value is to be used, statistical methods can be used to estimate the "best" single value for the parameter. Through Monte Carlo modeling, it is possible to use many parameter values for a single parameter, if needed.
By Manual Fit	One of the most common approaches to setting parameter values is to tak one or more sets of input and output data from the real system and then make many runs of the model, iteratively adjusting parameter values until "good" fit is achieved. This implies that the modeler has a firm idea of w constitutes a good fit. Taken to extremes, calibrating a model by fit treats the parameters as "fudge factors" to help make the model "fit" the real dat
Regression and Automated Fit	Regression is a more mathematically based approach to setting parameter values "by fit." In regression, varying parameter values optimizes an objective function (defining good fit). Common linear regression is the typical objective where the parameters of the model are optimized to find set of parameter values with the minimum sum of squared error. Where great amounts of input and output data are available for the real system, a the model equations are amenable to optimization, regression methods ca often yield statistics on the model's likely error and other quantitative estimates of goodness of fit. More sophisticated optimal parameter estimation techniques also are available (Beck 19893b).

In reality, several of the above methods are usually used to set values for model parameters. It is often something of an art.

From a different perspective, model calibration also is a form of model testing. If the model cannot be made to reasonably simulate known field observations by direct and reasonable modification of calibration parameters, then the model has in some way been tested and found to be empirically inadequate. Importantly, much can be learned from such failures, which are common in modeling. The ways that a model fails to "fit" a calibration data set also can be instructive in re-formulating the model by helping to identify specific processes or conditions that the model represents poorly (Beck 1985).

If a model can be adequately calibrated, additional testing, in the form of verification and validation is desirable. However, if the number of adjustable model parameters is large relative to the size of the calibration data set, then a "good fit" is often meaningless, since many sets of parameter values would likely give reasonable agreement with the small calibration data set. Large models with many adjustable parameters, typically require much larger calibration data sets.

Step 6. Model Verification.

Model verification can consist of several techniques that provide some test of the adequacy or reasonableness of the model for a particular purpose. Sometimes, model verification is defined as assessing if the model "behaves in the way the model builder wanted it to behave" (Beck 1983a; Gass 1983). Model verification and other model testing techniques are summarized in Table 3. Several such tests are typically employed, with specific tests applied to test particular model components in addition to testing overall model behavior.

Table 3. Methods for Model Testing (Verification and Validation)

Method	**Summary**
1. Sign Test	Do changes in model inputs lead to changes in model outputs in the "right" direction?
2. Ordinal Test	Do sequential changes in input values lead to output changes that are consistently in the "right" direction?
3. Sensitivity Analysis	Do changes in input and parameter values lead to "reasonable" changes in output values, both in magnitu and direction of change?
4. Turing Test	Can an "expert" in the subject of the simulation distinguish between the model's behavior and the behavior of the real system? Is model behavior "reasonable" to experts?
5. Comparison with Analytical Solutions	A test for numerical behavior, where rigorous analytic solutions exist for simple applications of the model, do analytical and simulation results agree?
6. Reproducibility or Comparison with Other Models	Do other studies and models find results similar to thos found by the model in question? For selected model components, do model results agree with hand calculations?

D-010094

7. Statistical Analysis	How much variation in the calibration data can be explained by the model? What is the statistical significance of the calibration of the model?
8. Independent Testing of Model Components	Confidence in the whole model is improved by testing individual model components.
9. Independent Calibration and Validation	Using separate data sets to calibrate and test the model, how well does the calibrated model estimate outputs fo the test data set? If several test data sets are available, what do these tests imply for the conditions that limit t model's effectiveness?
10. Deductive Proof	Can the model, or important parts of the model, be derived from fundamental information (e.g., conservati of mass, momentum, energy, and geometry)? Is the lo of the model correct and correctly implemented?

Many of the first four tests, particularly the Turing test, can be aided through some form of data display to aid the user and experts in evaluating the "reasonableness" of a large quantity of model results and the behavior of overall model results and the results of model components. A particular approach for the first four tests is "degenerate testing" (Sargent 1988), where model inputs are skewed to attempt to create degenerate model behavior, either for extreme cases under which actual system behavior is known (droughts drying reservoirs) or induce numerical or other logical degeneracy in the model's computations (e.g., large transients in dynamic models).

Sensitivity analysis is a common component in model development, used principally to test the reasonableness of model behavior, a form of model verification and to assess if particular components of the model need to be represented in more detail, or can be suitably represented with less detail. In this second function, if a sensitivity analysis shows that model outputs are insensitive to a particular parameter in a subprocess, then perhaps the representation of this sub-process can be simplified reasonably. Thus, some sub-processes within a system may be "parameterized," or represented by a single constant parameter. Conversely, if a model cannot be made to "fit" reasonable observed data without making a particular "parameter" vary in time or space, then perhaps a more detailed representation is needed of the parameter represented by that process. It is common for model verification to lead to improvements in model formulation.

Step 7. Model Validation.

The term "validation" is used variously in the computer modeling literature (Beck 1983a, b; Gass 1983). Here, model validation is the testing of a calibrated model by comparing model results with one or more sets of independent field or laboratory data. The intent is to provide an independent field test of the model, preferably under a variety of field conditions (such as wet and dry years). This is the highest scientific hypothesis testing form of model test. In terms of strength and rigor, it is superseded only by deductive proof from first principles.

The comparison of model results and field data for model validation is often not a simple exercise, but requires some consideration of which comparative statistics are appropriate for the particular objectives of the model. Comparative statistics could include (ASTM 1993):

comparative time-series of results (as tables or graphs) for specified locations, comparisons of maximum results (such as flood peaks), comparisons of duration above a water quality standard, or common statistical comparisons such as root mean squared error (RMSE), average absolute value of error, various types of correlation statistics, or statistical tests of the probability that model result distributions differ from the distribution of field data (which themselves may contain measurement errors).

Model validation is almost always difficult, requiring a large amount of independent high quality data. There are some problems for which model validation is prohibitively difficult, impossible, or irrelevant. An example of where validation is irrelevant is a water use or population long-term forecasting model. By the time enough future data is accumulated to verify the model, the forecasting use of the model is likely to have become mute. Models of complex processes, such as non-point source pollution or some complex operations problems also are difficult to verify, due to the difficulty of collecting spatially disaggregated data on a dynamic basis. Sediment transport models often are difficult to validate (as well as calibrate) because field data often are as prone to error as model results, making it difficult to compare model results and data (McAnally 1989). Often some sort of data validation is a desirable prelude to model validation. Where it is impossible practically to validate model results, the model may still have considerable use, although its detailed and quantitative results should not be viewed with the same confidence as results which closely correspond to accurate field data under a wide variety of conditions.

Where models are used for management, validation is always problematic, since the model is intended to examine system behavior under circumstances for which validation data is inherently unavailable, conditions which are significantly different from present conditions (Gass 1983; Thomann 1987).

Gass (1983) presents a broader view of model validation, including evaluation of the "face validity" of a model; is the model and its behavior reasonable to those with field experience with the system? This is much like the Turing, sensitivity, sign, and ordinal tests discussed in Table 3. In a sense, these are tests of the model's ability to simulate behavior seen in the real world.

Step 8. Documentation of Model.

> "The purpose of the model report is to communicate findings, to document the procedure and assumptions inherent in the study, and to provide general information for peer review. The report should be a complete document allowing reviewers and decision-makers to formulate their own opinions as to the credibility of the model." ASTM (1995a)

The three major forms of computer model documentation are as follows:

a) The computer interface with the user;
b) Comment statements in the source code; and
c) A manual or text, often in the form of user's and reference manuals.

In practice, documentation usually involves combinations of all these forms in varying amounts.

Although documentation via the model's user interface seems attractive, almost all models require more detailed documentation in the form of a separate simulation manual. The simulation manual is almost always the ultimate and authoritative form of documentation.

Comment statements in the source code (or notes in spreadsheets) are useful, but usually are only suitable for those who must dig into the model code and make changes. In essence, comment statements are directed to model programmers rather than model users.

The major documentation effort is the creation of user's and reference manuals. This is the text that most users will refer to when setting up to run a model, preparing data, and interpreting results. These manuals should describe the following (ASTM 1992, 1995b; Gass 1984):

- The particular objectives of the model and its range of applicability;
- The types of data required and the computer capability needed;
- The conceptual approach of the model;
- The mathematical formulation used in the model, and the limitations of this formulation;
- The numerical solution algorithm, including the limitations of this solution method;
- The calibration of the model and its performance in various verification and/or validation tests;

In addition, these manuals should include the following elements:

- Instructions for the user on how to run the model;
- Instructions for preparation of any required data files, including numerical size limits in this version of the model;
- A series of test cases that demonstrate the performance of the model;
- An example that leads the user through all steps in executing the model; and
- A set of references that allow the user to follow-up on particular aspects of the model.

Model documentation should be written clearly and precisely, with little use of jargon. The objective is to aid the user of the model and aid users of model results in interpreting and making use of model results. The model should not be a black box.

Because models are seldom fixed for long periods, the writer should establish a system of tracking version numbers. Whenever possible, the model should be structured at the beginning so that future updates and developments are easily understood by existing users and do not require re-entry of the input in a new format.

Step 9. Update and Support of Model.

The overall purpose of having an explicit process of model development is to increase the likelihood and degree that a model will serve the purposes for modeling discussed in Chapter 2. Just as we are more certain of the serviceability of a bridge if it is constructed from a well-analyzed and field-tested design and whose construction has been subject to inspection and component testing, a model that is methodically developed and implemented is far more likely to provide good service.

D-010098

Modeling Errors (Under Development)

Sources of Model Errors

Uncertainty errors in the appreciation of the system to be modeled

In solution attempts to problems associated with water management, one has to deal with systems that include two major, and quite distinct, types of components: the natural ones and the man-made ones. The latter ones are usually fairly well known because they were designed with specific criteria (one knows e.g. the dimensions of a spillway and its rating curve for discharge versus elevation or those of a concrete canal that conveys water from one part of a state to the other). On the other hand nature does not tell us what Manning' s roughness coefficient is for the innumerable heterogeneous segments of the rivers that crisscrosses the plains or valleys of the system, specially under extreme conditions of flood with overflowing banks, etc. Thus one must accept the fact that the system under modeling will always be described in a less than perfect way as to its physical static or dynamic characteristics. This introduces a fundamental error that cannot be circumvented and will always be present. The question that must be addressed is nevertheless, how to reach decisions in spite of the uncertainties associated with our comprehension of the system?

Conceptual errors in the description of the system to be modeled

Having understood that reality is too complex to be modeled perfectly, one must develop a schematic view of the system and its behavior. One needs to superimpose on reality our view of it and, naturally, a highly simplified one at that. In this process of conceptualization and simplification, an excellent understanding of the behavior of the natural system is needed in order to separate what is important and essential from what is secondary or tertiary. Conceptualization is required regarding both the static description of the system and its dynamic characteristics. In the static characteristics category one finds the geometrical and topographic description of say a watershed. How long is the main river? Can we approximate it by a succession of highly straightened segments with sharp turns at the junctions or must we subdivide the river into a very large number of reaches to accommodate changes in width, slope, roughness, direction, etc? Comes the hard question: how do we know the error committed by not carrying out the greatest level of refinement in the description? How do we test the model conception error resulting from a given level of coarseness in the description of such factors?
In the dynamic category one finds the description of the physical laws that govern the processes. It is fair to say that usually we know these laws at the molecular, microscopic or at the punctual scale in the continuum mechanics sense. At the molecular scale we know Henry' s law; at the microscopic scale we know Fick's law; at the punctual scale (or column scale, say cross-section of size 1 to 40 square centimeters, 20 cm deep, the Darcy scale) we know Darcy' s law. In water management we are not interested in these scales. We have to deal with thousands of square miles. To calculate infiltration in a basin as a result of rainfall events shall we then subdivide the top soil layer into billions of 25 cm2 cross-section , 20 cm deep, soil columns because we know Darcy' s law at that scale?

Hardly practical though some merchants (either peddling computers or watershed models) would like us to believe so!

At this point the infiltration process has to be conceptualized. There are several ways to proceed. A common way is to simply resort to analogy and schematically assimilate the topsoil to a lumped "reservoir" (or a few such reservoirs) with a variety of spillways and conduits. These spillways and conduits represent the transmission and retention characteristics of the soil at the area scale of 1 to 100 square kilometers and for depths from 1 to 10 meters. Another approach is to extend to a much larger scale say a satisfactory column law of infiltration (for example Horton' s or Green and Ampt' s), which depends on physically based parameters such as saturated hydraulic conductivity of the soil, and use in the formulas effective values of the parameters. Finally a third approach starts from the punctual laws and derives, more or less approximately, by multiple integration in space, time, expectation and process sense, a law valid at the larger scale of a parcel (1 to 10 square meters), hillslope (10 to 1000 square meters) or watershed (a few square kilometers). The problem with that approach is that to carry the integration one must conceptualize the laws of chance that underlie the spatial distribution of the soil parameters and the type and degree of connectivity between the columns or parcels that make up the hillslope or the watershed. We shall not here discuss the pros and cons of these distinct approaches; we want only to point out that all may potentially introduce serious model errors by failing to represent adequately the processes involved at the practical decision scale. Finally it is quite conceivable that some conceptual models commit the "sin of omission", i.e. do not account for a significant primary phenomenon. For example when studying infiltration into a layered soil, one must account for
the fact that some layers may be unsaturated and transmit by far less water than if they were assumed to be saturated. In this case the influence of capillarity was omitted.

Model equation errors

Starting from an analysis of the system and a conceptualization of its static and dynamic characteristics, one proceeds to express that knowledge in mathematical symbolism, leading to a system of equations and logical statements. The equations can be written in differential, integral and/or algebraic forms. Naturally if the conceptualization was severely in error or omitted significant processes, the mathematical formalism will not correct the analysis. Thus we shall here only discuss the additional errors that might be introduced at this level. If the form of the equations was differential in nature, again they only apply at the punctual scale. They have to be integrated. If it is done analytically then no error is introduced save for the possible wrong choice of boundary conditions! For example treating a river in hydraulic connection with an alluvial aquifer as a constant head boundary amounts to providing an inexhaustible source of recharge for the aquifer. This is acceptable if the river is the Mississippi but not the South Platte in Colorado.

Parameter estimation errors

In rare occasions the physical parameters appearing in the equations can be directly measured. For example there is no major problem in measuring the width of a river. However in almost all situations the parameters have to be estimated indirectly through a calibration procedure. The difficulty is that the estimation is always circumstantial and conditional. Pumping test procedures do not lead directly to transmissivity. From the measured drawdowns and the pumping rates in the pumped well one infers through an " inference model" (typically an analytical solution), which is based on a lot of assumptions) what the value of a uniform transmissivity would have to be so that the observations and the calculations by the inference model match in some "best" sense. Errors enter in the model because the assumptions of the inference model may be unreasonable for the given aquifer or because the selected criterion for the best match is not appropriate or because the match is fortuitous and only applicable under the conditions of the pumping test. Such estimations are usually conditional and circumstantial. For example the estimated parameters of an infiltration model may give a good match in the prediction of runoff given rainfall because all the events used for calibration displayed a single clearly marked peak but would be totally inapplicable if the event displayed several peaks. How does one quantify the error induced by parameters estimated for a certain type of conditions when one wishes to simulate different conditions? Typically the reason why models are built is precisely because one is not satisfied with current management. Thus future operations will often lead to events drawn from a different population than the one used in the calibration.

Input and data errors

Precipitation is measured at only a few points in space and what is needed is a continuous description of it. This is again something one has to live with as measurements, even if done without errors, will never cover the entire domain under investigation. Some time the temporal distribution of rainfall will have a significant one. In groundwater the Theiss solution is a favorite benchmark. However when the comparisons are made time steps and grid spacing are extremely small and the test results look good but in practice the time steps and grid sizes used are orders of magnitude larger than the ones used in these tests. What is the real truncation error for these large time and grid increments?

Confounding of errors

The most frustrating problem in trying to assess the magnitude of one sort of errors versus another kind is that they end up being confounded. For example if a finite difference model with coarse space and time increments is used to calibrate transmissivities based on observations, the parameters are conditional on the spacing. The same groundwater model used with a very fine grid (ignoring for now the problem of interpolation) may end up giving bad predictions because a wrong model with the wrong type of parameters when calibrated on good data will perform reasonably well, but it cannot be used reliably under different conditions than those for the calibration.

Interpretation errors

It is very easy specially when using automatic calibration procedures to come to a decent match but for the wrong reasons. Different models describing different mechanisms may lead to similar decent matches with proper adjustment of their parameters. However it is possible that with the calibration events the trigger for a completely different path in the computational sequence may never have been pulled because the necessary threshold is never reached. Thus a good understanding of the detailed workings of a model is needed for the interpretation of the results.

Sources of Modeling Errors

We use the term "modeling" to refer to the action of using a model for the purpose of conducting studies of water management.

"Wrong choice of model" error

The first task for a water manager is to choose a model that is appropriate for the purpose of the investigation to be performed. A wise user should be somewhat leery of using a model for rainfall-runoff that was developed in England and tested only there, for applications in Saudi Arabia. Either the basis for the model is thoroughly scrutinized and deemed applicable or a priori another model is selected. Naturally to scrutinize the model requires that a thorough documentation for the model be available. A poorly documented model should not be used even if it claims it can do the job, or you may suffer from a "trusting" error.

Wrong choice of increment error

This has been discussed in sections 1.4 and 1.6.

Wrong calibration and verification procedures

At some stage in the study process it will be necessary to calibrate a lot of parameters. First one must understand clearly the structure of the model to proceed with calibration. Let us say one has 20 years of daily data of rainfall and runoff for a catchment. Some parameters affect mass balance, some affects the medium term dynamics and some the short term dynamics. There will be evapotranspiration parameters, aquifer parameters, soil parameters and river parameters to calibrate. One should not attempt to calibrate all these parameters jointly on the 20 years of daily runoff. Rainfall and evapotranspiration will affect the long-term mass balance for runoff. Except for obvious transcription errors on rainfall these data will be taken at face value. Thus to establish a 20 years mass balance for the system one needs to adjust the coefficients involved in the calculation of evapotranspiration. If data of potential evapotranspiration were available derived from pan evaporation, typically a coefficient is adjusted to account for the variation between theoretical potential evaporation and the actual potential evaporation for the various watersheds in the basin. This coefficient is adjusted to guaranty a perfect mass balance for the 20 years of record, i.e. cumulative volume of runoff as calculated is the same as the observed one. The time scale for that calibration is 20 years. Next one can look at dry weather seasons. During these periods, over several months usually, runoff is driven by the parameters that condition aquifer recharge. One thus calibrates the recharge parameters on the characteristic shape of the recession curves of runoff. The time scale for calibration is now one or a few months. Next one looks at volumes of flood events for well characterized "single rainfall - single discharge peak events"

D-010102

to calibrate the parameters that control infiltration and thus excess rainfall and runoff. The time scale now is several days to a couple of weeks. Next one looks at the volumes under, and to a lesser degree the shape of, the discharge hydrographs under conditions of "double rainfall - double discharge peak events" with 1/2 day to couple of days separation between the rainfall events in order to calibrate the parameters that affect redistribution of moisture in the soil top two layers. The time scale is now a couple of days. Finally one looks at daily values during the flood events to estimate the parameters that control propagation and attenuation in the flood hydrographs. By now all parameters have been calibrated but unfortunately there is some dependence between the parameters so that the calibration steps must be repeated again from the large time scale down to the smallest scale at which the runoff data are known. This discussion illustrates that it is not possible to calibrate intelligently and reliably a model without both a good understanding of the phenomena and of the specific model structure to represent them. An error in modeling, which is not easily quantifiable, is the "lack-of-knowledge-induced" error. Now one would not use all the data to calibrate the parameters. One would select some of the years for calibration and some for verification. It is wise to select for calibration the years that did not exhibit extreme behaviors. The purpose of modeling studies is often to extrapolate to situations that have not been encountered through the historical record. Thus this partition in the record will demonstrate or not the ability of the model to extrapolate and will provide a quantification of the errors that are likely to be encountered for the more extreme situations. Naturally for actual use of the model in the future one now recalibrates the model with all the years of record, but one has a conservative basis for estimation of errors for the future.

Delimiting Significance of Errors (sensitivity analysis)

One needs to quantify the intrinsic model errors, i.e. the typical errors say in predicting runoff (if the model' s basic function is to predict runoff) as a function of seasons(dormant versus growing season) or flow conditions (e.g. during floods or recessions). One needs to study the impact of an error in estimation of certain parameters on certain quantities of interest such as instantaneous discharges or seasonal cumulative values. Is it important to know the exact value of transmissivity in the aquifer to predict peak discharge during a flood? In that case the answer would be no (practically) all the time and one could dispense with the exercise! Next is to study the impact of an error in certain parameters on a management decision. It is quite conceivable that a model that is not very accurate in predicting hydrograph shapes during flood events could be perfectly acceptable to size a new reservoir with interannual capacity. Whereas the intrinsic error can be secured once and for all by running the model, the "derived" errors for a particular study will depend on the objective of the study and an investigation is required for each individual study.

Characteristics of Useful Models

A model is useful, despite being wrong, when its quantified intrinsic error is compatible with the acceptable accuracy of a management decision made on the basis of use of the model. For example if a relatively large error in prediction of runoff during dry weather seasons leads nevertheless to a relatively narrow distribution in the needed size of a low flow augmentation dam, then it is useful. If the reverse holds the model in its present

D-010103

form is not useful. The intrinsic error must be reduced by a better calibration or by a change in structure or a combination. Even if the intrinsic error results in unacceptable level of error for the management decisions deduced from the use of the model, the model may have utility in a relative sense if one wishes to compare the effectiveness of different strategies. However one must be sure that the model incorporates properly the factors that condition the different responses between strategies. For example, if one is concerned about the effect of a pumping well near a river on a downstream surface water right holder, one must be sure that the groundwater model does not treat the river as a constant head boundary. In addition, it is important that the model account for the dynamic flow propagation and for the associated fluctuating river stage.

Model Validation Types Development (Under Development)

Evaluation of Mathematical Models

This section provides procedures and criteria for development, modification, and use of mathematical models. It does not specify models themselves, but it establishes minimum criteria for distinguishing acceptable models from those that may be incomplete, untested, or inappropriate for intended uses. This section permits the orderly evaluation of models for their completeness and suitability in specified uses. (ASTM, 1992).

Numerous pressures have led to the us of models in varying stages of completion, documentation, public availability, testing, and evaluation. Often models are used without any foreknowledge of the confidence that can be placed in their predictive capabilities. Models are sometimes used in ways that violate the assumptions and boundary conditions that are built into them. Such deficiencies and differences in models have led to unnecessary conflicts among users (ASTM, 1992).

Model Validation

Validation is the process of comparing model results to historical data. A model cannot completely duplicate historical data under all conditions for two reasons: (1) models are just mathematical representation of reality and (2) historical data contains problems with accuracy, precision, and completeness. Thus, validity is a matter of degree: it depends on the information available and is subject to the requirements established by the decision-maker. Despite this necessary level of subjectivity, models should not be used for assessments without examination of their validity (ASTM, 1992).

Modelers often state that their model "validates or verifies reasonably well" and then they point to a few stations and parameters that show "good" fit with historic data. However, to the decision-maker, which is usually the end model user, this general information is not usually very useful. To overcome this deficiency, the Forum has suggested a qualitative (and possibly quantitation) rating of the agreement between model predictions and historical data. What constitutes each level of validity should be clearly defined so that the reviewer can make an informed assessment. The table below outlines validation criteria that may be used by model users as a standard yardstick to evaluate models and their appropriate uses (that is, prevent use outside the limits of the assumptions used to formulate the model).

Table 4

Validation Type Criteria

Validation Type	Description	Criteria
I	Excellent	To be determined
II	Good	"
III	Fair	"
IV	Poor	"
V	Unacceptable	"

Public Access to Models

Computer models should be readily available to all users for independent evaluation prior to formal use as a decision-making tool. All model developers should archive an electronic copy of model in a manner that is accessible via the Internet. The Forum webpage will be enhanced to provide links to each organization's model webpage, if available. If necessary, the Forum will provide server space for storage of these models. Each log should include the model developer's name, the (executable and/or source) code, user's and reference manuals.

5. Use of Modeling in Planning Studies

Formulation and Evaluation of Alternatives Plans

An alternative plan consists of a system of structural and/or nonstructural measures, strategies, or programs formulated to alleviate specific problems associated with water-related resources in a particular planning area. Alternative plans should be formulated in a systematic manner to insure that all reasonable alternatives are evaluated. Probably the most common pitfall in planning is the failure to consider all alternatives, especially simple, yet non-traditional, alternatives. This occurs because the time constraints imposed on the project or apprehension of being ridiculed by management or peers.

The impact of the alternative plan is the difference between with-plan and without-plan conditions for each category of effects. Evaluation of alternative plans should be based on the most likely conditions expected to exist with and without the plan. The without-plan condition is the condition expected to prevail if no action is taken. The with-plan condition is the condition expected to prevail with a particular plan assumed to be in effect. The forecasts of with- and without-plan conditions should use "existing conditions" as the baseline.

Environmental impact analysis is an integral part of planning. The intent of California's CEQA and the federal NEPA process is to make Environmental Impact Reports (EIR) and/or Environmental Impact Statements (EIS) a decision-aiding document rather than the primary decision-making report. The environmental document should be integrated within the broader plan formulation/evaluation steps of the typical project planning process (Stakhiv, 1989)

An EIR is a detailed informational document that analyzes a project's significant effects and identifies mitigation measure and reasonable alternatives (CEQA Guidelines Secs. 15121(a), 15362). An EIR must describe the existing environmental setting from local and regional perspectives. When a proposed project is compared to an adopted plan, the analysis must examine existing physical conditions (CEQA Guidelines Sec. 15125). The EIR must always analyze the no-project alternative. The no-project alternative must describe maintenance of existing environmental conditions as a baseline for comparing the impacts of the alternatives (Dusek v. Redevelopment Agency (1986) 173 Cal.App.3d 1029) . For general plan amendments, however, it may be appropriate to analyze two no-project scenarios: maintenance of existing environmental conditions and future buildout under the existing general plan. These two approaches were used in the development of the draft State Water Resources Control Board (SWRCB) Decision 1630 (SWRCB, 1992). Draft Decision 1630 used the average annual historical export rate from 1984 through 1989 to represent existing conditions for all beneficial uses of Bay-Delta waters and estimated future demand for water supply planning purposes.

Decision 1630 stated:

> " Current estimated demand does not accurately predict the export rate that represents existing physical conditions, because (1) the estuarine

> ecosystem has never experienced the hydrological conditions that would exist if the current estimated demand were satisfied … (SWRCB, 1992)." "While exports will be less than would be expected in the future under (SWRCB) D-1485, the proper base for comparison to determine environmental effects is actual current conditions. (SWRCB, 1992)"

The 1984 through 1989 period is the most recent period before the 1989-1992 drought reduced exports and includes the largest export to date. The SWRCB concluded that the historical rate is recent enough to approximate existing physical conditions.

Similar issues have arose in developing the "environmental baseline" for the CVPIA use of the 800,000 acre-feet of "B2" water and recent SWRCB Bay-Delta Plans.

Documentation of Study Assumption

Planning studies should be documented in a clear, concise manner that explains the basic assumptions and decisions that were made and the reasons for them. An inventory should be made to determine the physical, chemical, and biological resource conditions. This inventory should describe the existing conditions and should be the baseline for forecasting with- and without-plan conditions.

Documentation of Modeling Study Results

Due to the importance of modeling to stakeholders and decision-makers, model studies (and development) need to be documented and archived to ensure quality assurance. The American Society for Testing and Materials (ASTM) has developed a framework for documenting and archiving a groundwater flow model application that can be tailored for Forum use (ASTM, 1995b).

Model documentation includes written and graphical presentations of model assumptions and objectives, the conceptual model, code description, model construction, model calibration, predictive simulations, and conclusions. Model archival refers to a file or set of files that contains logs of the calibration, sensitivity and predictive simulations, supplemental calculations, model documentation, a copy of the model source code(s) or executable files(s) used, or both, and input and output data sets for significant model simulations.

A model archive should consist of sufficient information generated during the modeling effort that a third party could adequately perform a post-modeling audit and such that future reuse of the model is possible. Table 1, which is reproduced electronically on the Forum webpage, shows the recommended simulation log that should be used to archive each significant model simulation. All model users should archive an electronic copy of the simulation log, including the (executable and/or source) code in a manner that is accessible via the Internet. The Forum webpage will be enhanced to provide links to each organization's simulation log webpage, if available. If not, the Forum will provide server space for these logs. Each log should include the modeler's name, simulation date, project name/number, simulation number, the code used (and version), the purpose of the run, the input file names, comments on the input data, the output file names, and comments on the results.

Table 5

MODEL SIMULATION LOG

By: ______________________________ Date: ______________

E-mail: ______________________________ Phone: (____)__________

Project Title and No.: ______________________________

Simulation Title and No.: ______________________________

Code Used/Version No.: ______________________________

Purpose of Simulation: ______________________________

____

____

Names of Input Files: ______________________________

____

Comments on Input Data: ______________________________

____

Names of Output Files: ______________________________

____

Comments on Results: ______________________________

____

____

____

General Comments: ______________________________

____

D-010108

7. Conclusions (Under Development)

References

American Society for Testing and Materials (ASTM) (1995a), "Standard guide for subsurface flow and transport modeling," ASTM D 5880-95, ASTM, West Conshohocken, PA.

American Society for Testing and Materials (ASTM) (1995b), "Standard guide for documenting a ground-water flow model application," ASTM D 5718-95, ASTM, West Conshohocken, PA.

American Society for Testing and Materials (ASTM) (1993), "Standard guide for comparing ground-water flow model simulations to site-specific information," ASTM D 5490-93, ASTM, West Conshohocken, PA.

American Society for Testing and Materials (ASTM) (1992), "Standard practice for evaluating mathematical models for the environmental fate of chemicals," ASTM E 978-92, ASTM, West Conshohocken, PA.

Bay-Delta Modeling Forum (BDMF) (1997), "Bylaws of Bay-Delta Modeling Forum," BDMF, Richmond, CA.

Bay-Delta Modeling Forum (BDMF) (1995), "BDMF Brochure," BDMF, Richmond, CA.

Bay-Delta Modeling Forum (BDMF) (1996), "Peer Review Process Paper," BDMF, Richmond, CA.

Beck, M.B. (1985), *Water Quality Management: a review of the development and application of mathematical models*, Springer-Verlag, New York.

Beck, M.B. (1983a), "A procedure for modeling," in G.T. Orlob (ed.) *Mathematical Modeling of Water Quality: Streams, Lakes, and Reservoirs*, John Wiley & Sons, N.Y.

Beck, M.B. (1983b), "Sensitivity analysis, calibration, and validation," in G.T. Orlob (ed.) *Mathematical Modeling of Water Quality: Streams, Lakes, and Reservoirs*, John Wiley & Sons, N.Y.

California State Senate (1997), "Senate Bill 1320, Environmental Regulations," California State Senate, Sacramento, CA.

Gass, Saul I. (1984), "Decision-aiding models: validation, assessment and related issues for policy analysis," *Operations Research*, Vol. 31, No. 4, pp. 601-631.

Gass, Saul I. (1983), "Documenting a computer-based model," *Interfaces*, Vol. 14, No. 3, May-June, pp. 84-93.

Gass, Saul I. and Bruce W. Thompson (1980), "Guidelines for model evaluation: an abridged version of the U.S. General Accounting Office Exposure draft," *Operations Research*, Vol. 28, No. 2, March-April, pp. 431-439.

Holling, C.S. (ed.) (1978), *Adaptive Environmental Assessment and Management*, John Wiley and Sons, New York, 377 pp.

James, A. (1993), "Simulation," in A. James (ed.), *An Introduction to Water Quality Modeling*, Wiley, N.Y., pp. 1-17.

Jacoby, S.L.S. and J.S. Kowalik (1980), *Mathematical Modeling with Computers*, Prentice-Hall, Englewood Cliffs, NJ.

Klemes, Vit (1982), "Empirical and Causal Models in Hydrology," in National Research Council Geophysics Study Committee, *Scientific Basis of Water-Resource Management*, National Academy Press, Washington, D.C., pp. 95-104.

Law, A.M. and W.D. Kelton (1991), *Simulation Modeling Analysis*, McGraw-Hill Book Co., N.Y.

Lund, Jay R. and Richard N. Palmer (March 1998), "Water Resource System Modeling for Conflict Resolution," *Water Resources Update*, University Council on Water Resources, pp. 10-12.

McAnally, W.H. (1989), "Lessons from ten years experience in 2D sediment modeling," in S.Y. Young (ed.), *Sediment Transport Modeling*, ASCE, N.Y., pp. 350-355.

National Research Council (1990), "Ground Water Models: Scientific and Regulatory Applications," *National Academy Press*, Washington, DC.

Scavia, Donald and Steven C. Chapra (1977), "Comparison of an ecological model of Lake Ontario and Phosphorus loading models," *J. Fish. Res. Board Canada*, Vol. 34, pp. 286-290.

Sargent, Robert G. (1988), "A tutorial on validation and verification of simulation models," in M. Abrams, P. Haigh, J. Comfort (eds.), *Proceedings of the 1988 Winter Simulation Conference*.

Stakhiv, Eugene Z., "The Role of the EIS in Water Resources Planning," *Water Resources Planning and Management: Proceedings of the 16th Annual Conference*, American Society of Civil Engineers, 1989.

State Water Resource Control Board (SWRCB), "Draft Water Right Decision 1630," SWRCB, Sacramento, CA, 1992.

Thomann, Robert V. (1987), "Systems analysis in water quality management - A 25 year retrospect," in M.B. Beck, *Systems Analysis in Water Quality Management*, Pergamon Press, N.Y.

U.S. Water Resource Council (1983), *Economic and Environmental Principles and Guidelines for Water and Related Land Resources Implementation Studies*, U.S. Water Resource Council, pp. ??-??).

Werick and Whipple (1994).

Appendices

I. Glossary (Under Development)

II. Technical and Management Modeling Questions that Stakeholder and Decision-makers Should Ask?

By Josh Collins (SFEI) 12/4/97

I think the "protocols" that you are attempting to construct could be very worthwhile. For some time I have been suggesting that the modeling community should begin to help bridge the information or knowledge gap that exists (and is growing) between the modelers and the resource managers who in many cases fund and must interpret models. In this regard I would like to suggest that you consider including perhaps as an appendix a series of technical and management questions that a resource managers ought to be able to ask about a model to gain a basic understanding of its applicability. Illustrative answers might also be provided, perhaps as a little dialog. For example, a Resource Manager might be advised to ask:

- What are the boundary conditions for the model?
- What are the assumptions of the model?
- How do they relate to each other and to the uncertainties of the output?
- What are the uncertainties or statistical confidences in the output and how could those be narrowed?
- What is the history of development of the model, and where else has it or a similar approach been used?
- Etc.

I doubt any terribly technical or mathematical answers would be appreciated, but rather some translation of that into basic understanding of where the model comes from, what are the assumptions, what are the inputs and outputs, and what are the uncertainties of those. I expect you can draw upon your wealth of experience to fashion such a QA, perhaps with a little help from associates in science and/or management.

I am forwarding this email to the hydrogeomorphic advisory team (HAT) of the Bay Area Wetlands Goals Project because the HAT has expressed some interest in this topic. I can't say whether or not they agree with me, but I think they might be able to help with your efforts.